DOCUMENT RESUME

ED 109 224 TM 004 685

AUTHOR Hawkins, Robert O., Jr.; Stolurow, K. Ann Coleman
TITLE Application of the Randomized Response Technique.
PUB DATE 1 May 75
NOTE 9p.; Paper presented at the New England Educational Research Association Annual Meeting (Provincetown, Mass., May 1, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS College Students; Confidentiality; Higher Education; Measurement Techniques; *Probability; Research Problems; *Response Mode; Response Style (Tests); Sexuality; *Statistical Analysis; *Statistics; *Validity
IDENTIFIERS *Randominzed Response

ABSTRACT

The Randominzed Response Technique was used with 83 undergraduate students in an Introductory Statistics course to: (1) demonstrate a means by which information of a sensitive nature can be obtained in a confidential manner; and (2) illustrate to a group of somewhat skeptical students an application of statistics to an interesting, real-world problem. The technique requires that each participating student respond to one of two questions, without revealing the one to which he/she is responding. The first question is neutral (e.g., Is the last digit of your student ID odd?); the other sensitive (i.e., Have you masturbated within the past two weeks?). The student is asked to choose a two-digit number (00-99) from a table of random numbers. Then, all students with numbers 00-69 are told to answer the sensitive question, and all with 70 or above answer the neutral question. Thus, only the student knows the question to which she/he is responding. The procedure provides an interesting demonstration of the utility of a statistical procedure applied to a practical, health science problem. (Author/DEP)

# Application of the Randomized Response Technique*

Robert O. Hawkins, Jr. and K. Ann Coleman Stolurow

School of Allied Health Professions

Health Sciences Center

State University of New York at Stony Brook

Stony Brook, New York 11794





* Presented at the New England Educational Research Association Annual Meeting, Provincetown, Mass.: May 1, 1975

Many of us who have taught undergraduate statistics have been faced with preconceived notions on the part of students that statistics is a bore, a waste of time and irrelevant. The attitude seems to be, "Everyone knows that statisticians are liars." This is not helped by the somewhat embarrassing fact that the reporting of social research has sometimes resulted in widely publicized misinterpretations of statistical data. Students are rarely familiar with some of the more interesting applications of probability and statistics.

An issue unlikely to appear in the popular press is that of the reliability and validity of methods used to obtain confidential or sensitive information. The majority of students are familiar with the work of Kinsey and Masters and Johnson, for example, and many are probably aware of the sample and response biases inherent in this work. They are not aware however, that topics covered in an elementary probability and statistics course could provide a simple means of reducing error associated with the collection of such sensitive information.

Warner (1965) proposed a procedure which utilized a concept called the randomized response. Subjects are asked to respond "True" or "False" to one of a set of related statements, such as:

A. "I masturbate at least once a week."

B. "I do not masturbate at least once a week."

Subjects do not indicate which of the questions they are responding to, and questions are assigned on the basis of a randomization technique (e.g., a table of random numbers.) Thus, the experimenter is able to manipulate beforehand only the proportion of subjects likely to be assigned to each group and not the identity of those individuals assigned. A "True" response, therefore, does not incriminate a respondent, since the group to which s/he was randomly assigned is unknown. This procedure does require a certain amount of attentiveness on the part of the respondent, particularly in the case of a "False" response to a negative statement.

A similar model was later proposed (Greenberg, Abul-Ela, Simmons, & Horvitz, 1969), which was designed to improve upon the earlier technique by introducing an unrelated question. This model was demonstrated with students as subjects by Campbell and Joiner (1973), and requires that one of the questions asked be sensitive and the other neutral. For example:

A. Have you masturbated within the past two weeks?

B. Is the last digit of your student ID number odd?

In order to determine which question to respond to, the student is asked to choose a two-digit number (00-99) from a table of random numbers. After the numbers have been chosen, all students with numbers 00-69 are told to answer Question A, and all with 70 or above to answer Question B. Thus, no one but the student knows the question to which s/he is responding, and confidentiality is ensured.

Responses are then collected, and the total number of Yes answers tabulated.

Estimating the percentage of students who answered the sensitive question affirmatively ($\pi$) requires the substitution of the following information:

$$\lambda = P\pi + (1-P)\theta, \quad \text{where}$$

$\lambda$ = proportion of Yes responses

$P$ = probability of asking the sensitive question

$\theta$ = fraction of students with odd ID numbers

This technique was implemented in an introductory statistics course to meet the following objectives:

(1) demonstrate a means by which information of a sensitive nature can be obtained in a confidential manner

(2) stimulate interest in statistics by illustrating to a group of somewhat skeptical students an application of statistics to an interesting, real-world problem.

## METHOD

A total of 83 health science undergraduate students participated in the study. These students had already learned to use a random numbers table and had participated in discussions of simple and conditional probability. They were told that, with this knowledge, they would be able to estimate the percentage of the group engaged in some particular sexual behavior and that the topic in question was considered by researchers to be sensitive. The procedure was explained without specifying the sexual behavior topic. The response to the proposed exercise was one of general skepticism yet at the same time, high curiosity and an eagerness to participate.

Using a table of random numbers, all students chose a two digit number, after which the two questions were written on the board.

Question A: Have you masturbated within the last two weeks?

Question B: Is the last digit of your ID number odd?

The students were then directed to respond to one of the two questions depending on the random number which they had selected (00-69 answer question A, 70-99 answer question B), and to write only "yes" or "no" on the piece of paper which had been handed to them at the beginning of the exercise. The responses were then folded and passed to the front of the lecture hall, and the total number of Yes responses tabulated.

## RESULTS

Substituting the following:

$$\lambda = P\pi + (1-P)\Theta, \quad \text{where}$$

$\lambda$ = proportion of Yes responses (.434)

$P$ = probability of asking the sensitive question (.7)

$1-P$ = probability of asking the neutral question (.3)

$\Theta$ = proportion of students with odd ID's (.5)

and solving for $\pi$, (i.e., the percentage of students who had masturbated within the past two weeks), it was determined that 40.5 percent of the responses obtained were affirmative. The students then were asked if they wished to validate the findings. They were cautioned that it would mean that each one would have to honestly answer the sensitive question, but were assured that the responses would be kept anonymous. It was explained that they would use a second sheet of paper that was to be handed out by the instructor,

cover the paper with a book or other material so that those sitting near would not be able to see what was written, answer the question, fold the paper several times and pass the paper forward to be collected by the instructor. All students agreed to answer the sensitive question to see what the actual result would be in comparison to the result obtained by the randomized response technique. The paper was handed out, the students responded confidentially and the papers were passed to the front of the hall. A tally of yeses revealed 36% responding positively. Of course, there is no way to ensure the validity of the responses in this situation, but one suspects the evasive answer bias to be less potent than in the traditional mode.

## CONCLUSION

Once one overcomes the hurdle of skepticism and the logistics of getting a group of respondents to use a table of random numbers properly, the technique provides an interesting demonstration of the utility of a statistical procedure applied to a practical research problem. By judicious selection, the sensitive question used can generate interest on the part of students. Because the technique is based on simple probability theory and requires only the use of a random number table, it is easily understood by students in an introductory statistics course. Further, this technique is particularly useful given the climate of the times--the restrictions placed on research on human subjects, and the general distrust of social scientists on the part of the public.

This combination of factors, while increasing the utility of the technique,

also serves to inhibit the conduct of studies which could be done to demonstrate its validity. Studies of honesty (e.g., Hartshorne, May and Shuttleworth, 1930) provide a criterion for determining the validity of the responses, but given the guidelines for human subject research, it is doubtful that such studies would be approved in the near future. The technique does appear to be a useful one for gathering sensitive information, and in addition provides an interest-provoking learning experience for students in an undergraduate statistics course.

## References

Campbell, D., & Joiner, B.L. How to get the answer without being sure you've asked the question. The American Statistician, 1973, 27, 229-231.

Greenberg, B.G., Abul-Ela, A.A., Simmons, W.R., & Horvitz, D.G. The unrelated question randomized response model, theoretical framework. Journal of the American Statistical Association, 1969, 64, 520-539.

Hartshorne, H., May, M.A., & Shuttleworth, F.K. Studies in the organization of character. New York: MacMillan, 1930.

Warner, S.L. Randomized response: A survey technique for eliminating evasive answer bias. Journal of the American Statistical Association, 1965, 60, 63-69.